

- Reduction of maternal-infant transmission of human immunodeficiency virus type 1 with zidovudine treatment. *N Engl J Med* 1994;331:1173-80.
- 11 European Collaborative Study. Caesarian section and risk of vertical transmission of HIV-1 infection. *Lancet* 1994;343:1464-7.
 - 12 Dunn DT, Newell ML, ADES AE, Peckham CS. Risk of human immunodeficiency virus type 1 transmission through breastfeeding. *Lancet* 1992;340:585-8.
 - 13 Department of Health. *Guidelines for offering voluntary named HIV antibody testing to women receiving antenatal care*. London: DoH, 1992. (Guidance PL/CO(92)5, appendix 2.)
 - 14 Department of Health. *Guidelines for offering voluntary named HIV antibody testing to women receiving antenatal care*. London: DoH, 1994.
 - 15 Department of Health. *Changing childbirth*. London: HMSO, 1993. (Report of the Expert Maternity Group.)
 - 16 Flesch R. A new readability yardstick. *J Appl Psychol* 1948;32:221-35.
 - 17 Meadows J, Catlan J. HIV seroprevalence and antenatal clinics. *Lancet* 1992;339:622-3.
 - 18 Howard LC, Hawkins DA, Marwood R, Shanson DC, Gazzard BG. Transmission of human immunodeficiency virus by heterosexual contact with reference to antenatal screening. *Br J Obstet Gynaecol* 1989;96:135-9.
 - 19 Chrystie IL, Zander L, Tilzey A, Wolfe CDA, Kenney A, Banatvala JE. HIV testing in pregnancy: worthwhile? Can we afford it? *AIDS Care* 1995;7:135-42.
 - 20 Meadows J, Catalan J. Why do antenatal attenders decide to have the HIV antibody test? *Int J STD AIDS* 1994;5:400-4.
 - 21 Chrystie IL, Sumner D, Palmer S, Kenney A, Banatvala J. Selective screening of pregnant women for evidence of current hepatitis B infection: selective or universal? *Health Trends* 1992;24:13-5.
 - 22 Barbacci M, Repke JT, Chaisson RE. Routine prenatal screening for HIV infection. *Lancet* 1991;337:709-11.
 - 23 Cozen W, Mascola L, Enguidanos R, Bauch S, Giles M, Paxton P, et al. Screening for HIV and hepatitis B virus in Los Angeles County prenatal clinics: A demonstration project. *J Acquir Immune Defic Syndr* 1993;6:95-8.
 - 24 Delamothe T. HIV infection concentrated in London. *BMJ* 1995;310:213.

(Accepted 7 July 1995)

Improving oral examinations: selecting, training, and monitoring examiners for the MRCGP

Richard Wakeford, Lesley Southgate, Val Wass

Unless examiners are carefully selected, trained, and monitored, examinations may become hazardous. This is perhaps most true of oral or viva voce ("viva") examinations, which can generate marks unrelated to competence. To help other bodies to short circuit some years of experiment in connection with the oral component of the Royal College of General Practitioners' membership examination (MRCGP), this paper describes the selection, training, guidance, and monitoring arrangements that have been developed.

The oral or viva voce examination ("viva")—a general non-patient based encounter between a candidate and one or more examiners—has held an important place in medicine for centuries.¹ Tradition aside, it is used for its flexibility, its apparent fidelity (much medicine concerns oral encounters over issues of diagnosis and management), and its potential for testing higher order cognitive skills.

Unfortunately oral examinations are prone to many errors.² These include errors relating to halo effects (a judgment of one attribute influences judgments of others); errors of central tendency (judgments cluster in the middle); so called errors of logic (mistakes); a general tendency towards leniency; and errors of contrast (judgments of a candidate are influenced by impressions of preceding candidates). Oral examinations tend to test at a low taxonomic level—for example, factual knowledge rather than problem solving.³ Scores are related to irrelevant attributes of candidates such as appearance or confidence.⁴ Agreement between examiners is often poor.⁴ It is, moreover, difficult to establish in any formal way how valid an oral examination is.⁵

Largely abandoned in North America, oral examinations are still widely used in undergraduate and postgraduate examinations in the United Kingdom. In 1990, 19 out of 27 medical schools used vivas in their final qualifying examinations, 11 as a major assessment method.⁶

The membership examination of the Royal College of General Practitioners (MRCGP) also uses oral examinations. Hitherto, a practical clinical component has not been thought feasible as there are some 2000 candidates each year. Given the centrality of the consultation in general practice, an oral examination component has been regarded as appropriate, often being referred to informally as a "clinical by proxy."

Evidence suggests that oral examinations can be improved by the careful selection⁷ and training⁸ of

examiners. Much effort has been expended towards enhancing the reliability and validity of the MRCGP, especially by addressing the selection, training, and monitoring of examiners. We describe our approaches and make general recommendations from our experience.

MRCGP examination

The MRCGP examination currently comprises three written components—a multiple choice test, a modified essay question paper, and a critical reading question paper—and two half hour oral vivas, each conducted by two examiners. Major changes to the examination, planned for 1996, will not affect the vivas.

Poor performance on the aggregate of the written papers excludes some candidates from the oral examinations. A pass is achieved on the aggregate of the five marks. Statistical monitoring ensures that each component contributes equally to the total. The examination and its impact on candidates' learning behaviour have been described.^{9 10}

Selection of examiners

Examiners need knowledge and skills relating to their subject and towards participation in the design and conduct of the examination. The ability to conduct oral examinations effectively (and to participate in the planning of the written components) requires three further attributes: an approach to the practice of medicine and the delivery of health care that is within the limits of that acceptable to the examiners as a whole; effective interpersonal skills; and the ability to act as a productive member of a small team.

Examiners have consistent and contrasting marketing behaviours—for example, hawk or dove and restrained or theatrical.¹¹ Unless extreme (for instance, an examiner who gives the same mark to all candidates), such behaviours are containable, as they can be changed either statistically or by training.^{12 13} Behaviour that cannot readily be changed, however, is disagreeing with fellow examiners about what is a better answer and what is a worse answer—in other words, a low rank order correlation with colleagues' marks.

To ensure that examiners are of high quality the college requires that potential examiners must be members of the college and in active general practice. If they took the examination more than 10 years ago, they must retake it. They are required to undertake

The Old Schools,
Cambridge University,
Cambridge CB2 1TT
Richard Wakeford,
staff development officer

Medical Colleges of
St Bartholomew's and the
Royal London Hospitals,
London EC1M 6BQ
Lesley Southgate, professor
of general practice

Chislehurst, Kent
BR7 5AX
Val Wass, general practitioner

Correspondence to:
Mr Wakeford.

BMJ 1995;311:931-5

FIG 1—Examiners' grading sheet for oral component of MRCGP examination (revised in 1994)

two activities before they are assessed: answering a selection of recent written questions and marking 20 selected examination papers. They also sit half of a recent multiple choice paper.

The results of the potential examiners' marks on the written work and the estimates of how well their marking agrees with each other and that of the original examiners are analysed. The Examination Board's educational consultant advises whether any of these should give cause for concern.

On the day of assessment, potential examiners are asked to undertake two further tasks: (a) simulated group work on setting and marking written questions and (b) simulated oral examinations, acting in turn as examiner, candidate, and observer. Over a day potential examiners are thus observed undertaking interactive procedures that constitute the core of activities relating to the examination. Experienced examiners independently judge them in terms of their approach to general practice and their skills in interpersonal communication and teamwork. At the end of the day the performance of each potential examiner is reviewed at a meeting chaired by an independent person; the names of those about whom there is serious concern are noted.

When concern about a potential examiner emerges from both parts of this procedure he or she is not

invited to join the panel of examiners. The letter sent emphasises that examining skills and clinical skills are different and that rejection as an examiner is no criticism of the person's work as a doctor. (This is analogous to comparisons between the skills needed to interview patients and those needed to interview candidates for a job.¹⁴)

Examiners are first accepted for a probationary period of two years. Training takes place as described below, but after 18 months, after a routine video training session, a new examiner's performance is formally reviewed against a set of criteria derived from analyses of examiners' tasks (see box 1). A poorly performing examiner might be counselled to undertake additional, specified training or not to seek formal reappointment.

Examiners' tasks in the oral examination

The parts of an examination need to be defined in terms of their function and content. The function of the oral component of the MRCGP examination is to judge candidates' approach to practice, their decision making skills, and their justification for their decisions.

More difficult, though, when there is no detailed syllabus for the examination, is to define the boundaries of the content of the oral examinations. This is now

Box 1—Examiners' tasks in oral examinations

- To mark at least six different topics in each half hour viva
- To include adequate exploration of all agreed areas of competence specified for the examination—for example, diagnosis, therapeutics, communication skills, ethics
- To explore the candidate's approach to the practice of medicine, searching for coherence, rationality, and consistency
- To obtain justification of reported behaviours, approaches, opinions, and attitudes—for example, by reference to published work
- To attempt to link stated behaviours and approaches to performance
- To avoid topics included in current written examination papers, as appropriate
- To grade the candidate on each topic (to include recording each grade)
- To make and record an overall judgment, with weighting of individual topics
- To conduct the examination with respect for the candidate and fairly with regard to equal opportunities for all

undertaken by means of a modified Delphi technique, codifying the views of the panel of examiners as a whole. Such a study was first undertaken in 1985 and has been repeated in 1994.⁹ The attributes identified could be clustered under seven headings, which were used to create a grid (fig 1) to encourage examiners to balance their oral examination and constrain it within the agreed limits, and to be more systematic generally.

A final constraint is the so called high case specificity of performance in medicine generally: a doctor's performance in solving different problems may vary substantially as doctors are often good at some things and bad at others.¹⁵ It is vital that as many topics as possible (at least six) are covered in each viva. Such considerations have resulted in a specification for the tasks of examiners during an oral examination (box 1).

Problems and strategies for oral examiners

Discussions with examiners over 10 years and watching video recorded examinations have identified a variety of practical problems which examiners face.

O	Outstanding	A very rare candidate. Uniformly outstanding. Well read, coherent, rational, consistent, critical. Without being asked, justifies approaches, etc, by reference to published work.
E	Excellent	Extremely impressive candidate. Generally outstanding candidate but not so uniformly well informed.
G	Good	Generally impressive candidate. Well informed, coherent policies, fairly critical. Good decision making skills. Justifies majority of approaches well.
S	Satisfactory	A candidate characterised by a reassuring solidness rather than impressiveness. Able to justify only some approaches well, but most appear sensible. Adequate, not good decision making skills.
B	Bare pass	Examiner is only just comfortable with candidate's adequacy. Not much justification of approaches, but important ones are sensible. Decision making and other skills tested are just, on balance, acceptable.
N	Not very good	Questionable approaches, sometimes neither justifiable nor justified. Examiner is uncomfortable with candidate and his or her decision making skills, thinking him or her to be possibly risky in practice. Seems not to be good at applying basic knowledge.
U	Unsatisfactory	Approaches are often inconsistent and rarely justified. Candidate does not seem to be capable of passing the examination overall. Poor at applying knowledge.
P	Poor	Candidate clearly not passable, though slight evidence of ability. Generally incoherent approach to practice. No justification for specific approaches.
D	Dangerous	Candidate is worse than poor. Adopts such arbitrary approaches as to put patients at risk.

FIG 2—Grades in oral component of MRCGP examination

Little published work applies to oral examinations in higher or medical education, but a related subject is the selection interview, which forms a focus in occupational psychology research.¹⁶ A summary of problems derived from both sources is shown in box 2.

Reviewing the tasks of examiners in the light of these problems and with the benefit of experience has led us to identify strategies for planning oral examinations (box 3) and practical techniques to assist in their conduct (box 4).

Box 2—Problems for oral examiners

- Practical problems experienced in practice
 - Dysfunctional start to the oral
 - Difficulty in covering the ground fast enough
 - Problem candidates (for example, slow spoken, slow witted, or garrulous)
 - Losing control to the candidate
 - Candidate talking about unmarkable yet cognate issues—for example, training experiences
 - Coexaminer overrunning on a topic
 - An uncomfortable or dysfunctional end to the oral
 - Disagreeing with coexaminer about the overall grade
- Problems adduced from published work on selection interviews
 - First impressions will be overly influential on a final judgment
 - The appearance (attractiveness, particularly) of a candidate will influence the grade given
 - The contrast with previous candidates can affect an examiner's judgment—for example, after two poor candidates a moderate candidate may seem very good
 - Examiners will tend to treat preferentially people like themselves—for example, those holding similar values—and people they like
 - Examiners may be especially critical of faults in candidates which they know they also have
 - Examiners are trying to make global unidimensional judgments of people such as "good" or "poor"; in practice, most candidates will have good and bad aspects

Grading and marking

Two main difficulties confront the examiner when marking an examination: the varying attributes given to numbers under different marking systems that examiners may be used to—for example, 55 may be the pass mark under one system and 53 a good pass under another—and how to reach an overall mark from several component marks.

To avoid these problems and to encourage examiners to think about a candidate's performance in the oral examination as a whole, we have developed a grading scale based on simple epithets and more extended descriptions of these (fig 2). General guidance given about grading and marking is shown in box 5.

Training of examiners

Examiners for the MRCGP examination are trained by participating in preliminary practice vivas; by observing vivas; by regularly reviewing themselves on video and receiving feedback from others; and by attending an annual workshop.

When they are appointed, new examiners undertake and review practice vivas with volunteer trainees. Supernumerary examiners regularly observe oral examinations, and review them afterwards with the actual examiners. In this way, questions are refined and standards discussed in a way that is helpful to both the observer and the observed, with exchanges of ideas

and approaches. At the end of a morning of observing, a structured review session enables observers (and any visitors) to discuss issues with members of the royal college's oral development group.

Formal review of and feedback from videos is provided roughly every year to all examiners. Two pairs of examiners who see the same candidates are recorded on video during a morning examination session of six oral examinations. The afternoon is set aside for them to review the recordings with an educational consultant (a psychologist). There are two sessions. The agenda for the first is set by the consultant (RW), who identifies teaching points for each examiner and identifies them by means of excerpts from the video. The second session is devoted to a consideration of the vivas that caused the examiners the most concern or interest and which they wish to review. These sessions are guided by rules for feedback which ensure that it is supportive at the same time as being effective.

Each year, examiners are invited to attend a three day workshop, which serves various functions. In particular, it affords an opportunity to discuss with the panel as a whole possible developments and directions for change. Training can be targeted at specific groups of examiners. And new approaches can be practised by everyone, together, and difficulties collectively resolved.

Box 3—Strategies in planning vivas

- Plan to start a new topic every five minutes (at least six topics in a half hour oral)
- With coexaminer, plan an overall structure which samples from as many different topics and areas of competence as possible
- Have questions written on cards and classified by subject of competence being assessed. This facilitates planning and enables prompt action when things do not go to plan
- Have available a short term emergency question for use when things go wrong—for example, when coexaminer finishes unexpectedly or becomes indisposed

Box 4—Techniques during viva

- Spend a few moments initially putting a candidate at ease, shaking hands, and inviting a comment about transport or weather; this develops rapport and avoids initial dysfunction
- Introduce each topic and define its area. For example, "I'm going to ask you about juvenile onset diabetes, but I want to concentrate on issues of doctor-patient communication"
- Because you are limited to five minutes per topic, go to the core of the question quickly ("what I'm really getting at..."), using short questions and avoiding the verisimilitude of detailed scenarios, which waste time
- Avoid factual questions and unmarkable questions. Factual knowledge cannot be reliably tested in a viva and must be the focus of a written test. Unmarkable questions produce information on which you cannot make judgment—for example, about a candidate's previous colleagues
- If you plan to use props (letters, pictures, electrocardiograms, etc) make sure that their function is clear and that they enhance the testing process and do not waste time. (Our experience in the MRCGP examination is that they rarely add much and often waste time and confuse)
- There are often no clear cut right and wrong answers in medicine. Because of this, it may be helpful

Box 5—Grading and marking

- Each topic should be graded by reference to the list of grades and descriptions. If you are unhappy about a grade's accuracy, annotate it—for example with brackets. Make a note of reasons for giving a grade
- Use some questions regularly for calibration purposes. Note on your record card characteristics of answers from poor, average, and good candidates
- If you are giving a fairly high grade to a topic ask yourself what the candidate would have to have done better to get a better grade? We find that in this way, examiners may extend their use of the grading scale
- At the end of the viva review the list of grades given to each topic. Refer to the list of grades and descriptions: which fits the candidate best? Is an average obvious? If not, consider the firmness of each mark: does this help? Otherwise, are there any examination policies—for example, to err on the side of generosity or caution—which will help you?
- When considering your overall grade, review the list of hidden problems in box 2. Would these on balance be tending to push your mark inappropriately high or low?
- Beware of the common feeling that candidates improve towards the end of a viva and thus raising a grade. This feeling is more likely to reflect true variations in a candidate's ability among the topics discussed than the candidate's true ability
- Do not let your coexaminer browbeat you into changing your grade. Independent judgments are required. Unless it transpires that you have missed a catastrophic or brilliant answer, maintain your judgment

Discussion

The reliability of oral examinations can best be estimated from the extent of agreement between pairs of examiners. The correlation between the two examiners' judgments has steadily risen, and the percentage of grades from the two that are within one grade of each other is now 94%. This encouraging position is seen as resulting from careful selection,

to use a model when presenting a question of choice—for example, the "options, implications, decision, justification" model, asking: What are the options open to you now? What are the implications of each? What would you do? What is the justification for this decision?

- Plan tactics for difficult candidates—for example, a slow candidate may be encouraged non-verbally and with specific questions ("Gived me three advantages of..."), and such questions can be used to control an overbearing, bulldozing candidate. A garrulous candidate may be slowed by asking for clarification and interrupted and controlled with body language
- Poor candidates may need to be encouraged. But for legal reasons, avoid using terms that may be taken as a statement that they are doing well—for example, "that's good." Best, use non-verbal encouragement
- Arrange a code for communicating with your coexaminer if he or she overruns
- When you feel you can award a grade to a topic, do so and finish. The more topics covered in an oral examination the better
- When the bell goes, let the candidate finish his or her sentence before closing the examination. Otherwise, examiners may seem abrupt to the point of rudeness when stopping candidates at the bell

monitoring, and training of examiners, including defining the function of the oral examination within the overall examination and specifying its process.

Conducting an effective oral examination requires a great deal of commitment and effort. Without commitment and effort you are likely to generate something approaching random numbers. This should be noted by examining bodies who give equal weight to marks in written papers, clinical examinations, and vivas to obtain an overall mark. We believe that five key elements provide a defensible oral examination.

- Identifying the main tasks of examiners, and selecting examiners for these tasks
- Careful planning of each oral examination as a whole
- Contingency planning for difficult candidates
- Providing preliminary and ongoing training of a supportive nature, and ensuring the participation of all examiners in continuing discussions about the oral component and its development
- Monitoring the examiners and the examination overall, both statistically and within the training process.

RW is consultant to the Examination Board of Council of the Royal College of General Practitioners, LS is convenor of the Panel of Examiners, and VW is the convenor of the board's Oral Development Group. Much of the content of this paper has been generated in discussions with college examiners, and we thank them for their contributions, especially past and present members of the Oral Development

Group, Peter Tate and George Smerdon in particular. This article expresses our views and not those of the college.

- 1 Anderson J, Roberts FJ. *A new look at medical education*. London: Pitman Medical, 1965.
- 2 Guilford JP. *Psychometric methods*. New York: McGraw Hill, 1954.
- 3 Evans LR, Ingersoll RW, Smith EJ. The reliability, validity and taxonomic structure of the oral examination. *Journal of Medical Education* 1966;41: 651-7.
- 4 Thomas CS, Melliss G, Callender J, Crawshaw J, Ellis PM, Hall A, *et al*. The oral examination: a study of academic and non-academic factors. *Med Educ* 1993;27:433-9.
- 5 Norman GR, Muzzin LJ, Williams RG, Swanson DB. Simulations in health sciences education. *Journal of Instructional Development* 1985;8:11-7.
- 6 Wakeford R. *Report of the Survey of Undergraduate Medical Education 1990. Vol 1. General policies*. London: General Medical Council, 1992:53-4.
- 7 Newble DJ, Hoare J, Sheldrake PF. The selection and training of examiners for clinical examinations. *Med Educ* 1980;14:345-9.
- 8 DesMarchais JE, Jean P, Delorme P. Training in the art of asking questions at oral examinations. *Annals of the Royal College of Physicians and Surgeons of Canada* 1989;22:213-6.
- 9 Lockie C, ed. *Examination for membership of the Royal College of General Practitioners (MRCGP)*. London: Royal College of General Practitioners, 1990.
- 10 Wakeford R, Southgate L. Postgraduate medical education: modifying trainees' study approaches by changing the examination. *Teaching and Learning in Medicine* 1992;4:210-3.
- 11 Wakeford RE, Roberts S. Short answer questions in an undergraduate qualifying examination: a study of examiner variability. *Med Educ* 1984;18: 168-73.
- 12 DesMarchais JE, Jean P. Effects of examiner training on open-ended, higher taxonomic level questioning in oral certification examinations. *Teaching and Learning in Medicine* 1993;5:24-8.
- 13 Raymond MR, Webb LC, Houston WM. Correcting performance-rating errors in oral examinations. *Evaluation and the Health Professions* 1991;14: 100-22.
- 14 Pendleton D, Wakeford R. Interviews in the selection of partners, trainees and medical students. *Journal of the Royal College of General Practitioners* 1988;38:147-8.
- 15 Newble D, Jolly B, Wakeford R. *The certification and recertification of doctors: issues in the assessment of clinical competence*. Cambridge: Cambridge University Press, 1994:233-4.
- 16 Cook M. *Personnel selection and productivity*. Chichester: Wiley, 1988:45-64.

(Accepted 14 July 1995)

Letter from Cuba

Cuba: plenty of care, few condoms, no corruption

Hans Veeken

The health system in Cuba guarantees accessibility to the entire population, is free of charge, and covers the spectrum from vaccinations to sophisticated interventions. The results are impressive: Cuba's health figures are on a par with developed countries that have 20 times the budget. The country is experiencing a difficult period because of the collapse and loss of support from the Soviet Union; over 30 years' trade embargo by the United States; and the gradual change from a centrally planned economy towards more of a free market system. Shortages are experienced in every sector, and maintaining health care services at the current level is too expensive. Doctors and nurses continue to work towards the goal of health for all Cubans, even though their salaries are minimal. Signs of negligence or corruption, often seen in other socialist countries where incentives for output are lacking, are unknown. Topics such as family planning and AIDS deserve immediate attention.

"Cubans were the first '*medicos sin fronteras*,'" the government representative answers with a smile when we introduce ourselves at the Ministry of Health. "We have always exported doctors to places where they are needed; at one time there were 5000 Cuban doctors abroad. No, doctors we are not short of, but we could use some help with drugs and supplies. The recent economic crisis makes it difficult to buy them and health should remain a priority for our people. Our system guarantees total accessibility, is free of charge,

and covers the whole spectrum from vaccination to heart transplantation," he continues. "We started with 6000 doctors after the revolution; 3000 of them immediately left for the States. Now we have 60 000 doctors, one for each 200 inhabitants," he says proudly.

Care in Cuba

Each year around 4000 students start their medical training at 23 different universities. The Cuban health system consists of three tiers. The first tier is the front line care provided by family doctors. Since 1980 there has been a training for family doctors. Now 22 000 of them have been trained and they cover 90% of the population. Their main work is preventive: health promotion and offering basic curative care. The family doctors are backed up by 400 polyclinics, where specialists offer their services to about 30 000 people. The 263 hospitals form the third tier of the Cuban system.

The progress Cuba made in the field of health is impressive. Cuba's health system always got priority, as did education. Health indices are on a par with developed countries, yet the total budget of the country is a tenth to a twentieth of theirs. Life expectancy is 77 years and infant mortality a mere 9 per 1000 live births. The common diseases of poverty have been effectively combatted. Malaria has been eradicated from the island and Dengue fever successfully reduced by an immense campaign in which the whole population participated. Leading causes of

Médecins Sans Frontières,
PO Box 1001, 1001 EA
Amsterdam, Netherlands
Hans Veeken, public health
consultant

BMJ 1995;311:935-7